A Simulation Comparing Sampling Techniques for Estimating the Use of Preventive Care

Kris Moore, Dwight Stephens, Baylor University; Scott Burk of Scott and White Memorial Hospital
Dwight Stephens, Baylor University, 522 S. University Parks Dr. #903, Waco, TX 76706

Key Words: Systematic, Stratified, Stratified Systematic, Random sampling

Introduction

The National Committee for Quality Assurance (NCQA) is an organization that independently accredits managed care plans. The NCQA produces a set of information called the Health Plan Employer Data and Information Set (HEDIS) to help consumers, employers, and public purchasers compare health plans. Information is provided by measurements selected by the NCQA that attempts to evaluate performance and prevention components in health care (De Lafuente, 1996). Participating health plans sample claims and medical records in order to provide measurements of service utilization. The NCQA recommends random and systematic sampling methods for obtaining measurements, but the NCQA will also allow more complex sampling methods, including stratified and cluster sampling in order to produce measurements that are more efficient. Plans that use a more complex sampling method must show that the sampling approach is auditable and is without bias against a particular subject being chosen (HEDIS 3.0, 1998). In this paper, a simulation was performed to compare the effectiveness of random, systematic, stratified, and stratified systematic sampling plans in predicting effectiveness of care measures. The results of the simulation indicated that a stratified or a stratified systematic sampling plan could improve the quality of the data collected when the differences between strata is moderate to large and the sample size is small.

HEDIS

HEDIS was created by a panel that included public and private insurance purchasers as well as representatives of labor unions and health plans (De Lafuente, 1996). HEDIS 3.0, introduced in January 1997, contains seventy-one measures of clinical performance, procedure utilization, and patients' experiences and satisfaction with care (Thompson, et al., 1998). These measures cover eight major areas of interest: effectiveness of care, availability of care, satisfaction with care experience, health plan stability, use of services, cost of care, informed health care choices, and health plan descriptive information (Grimaldi, 1997). Health plans participate on a voluntary basis and submit their data electronically (Thompson, et al., 1998).

Information Consumers

In recent years, there has been an unprecedented demand for measures of quality in the health care industry. HEDIS was created to evaluate the value of managed care plans and to provide accountability for performance of plans (Nustad, 1997). The NCQA compiles and publishes information on HEDIS in the Quality Compass database which currently consists of data primarily used by commercial health plan purchasers, but there are efforts to collect information for public purchasers as well. For example, Healthcare Finance Administration (HCFA) is currently requiring HEDIS measures to be reported for plans contracting to provide care to Medicare beneficiaries and a number of states are also using the measures to monitor the quality of care provided by Medicaid managed care providers. Large companies have started to use the measures to provide incentive based contracts with care providers based on their performance (Thompson, et al., 1998). The Quality Compass provides information to private purchasers, state agencies, and HCFA about how health plans operate and meet member's needs (Grimaldi, 1997). Unfortunately, there is little information collected in any standardized way for traditional fee-for-service providers (Thompson, et al., 1998).

How is information obtained?

The NCQA produces a database of information on the quality of care provided by most of the capitated health care systems in the United States in the Quality Compass. The database is an effort to standardize and collect specific plan information and make it available to individuals and groups in the marketplace. The data provides the ability to compare performance attributes of health care plans across the managed care industry and a measure of public accountability for the level of service provided by these health plans. The NCQA compiles the data into its database and reviews the data for accuracy, and also twenty percent of the participating plans submit their reports to external auditing before submitting their data (Thompson, et al., 1998).

Limitations

There are limitations to the data collected for HEDIS. Participation of commercial health plans in HEDIS is voluntary. Some plans may not report measures because they think that the potential financial risk of comparative performance data is greater than any potential gain of participating, and some participating health plans may be tempted to misrepresent their data as a result of perceived financial risk. Because of this potential threat, the NCQA introduced standardized audit procedures in 1998. In addition, many health plans do not consistently document all care administered and may score artificially lower on the measures (Thompson, et al., 1998). Unfortunately, not all providers can provide complete and accurate HEDIS reports. HEDIS may require health care providers to change or standardize policies and procedures of service delivery in order to ensure reporting accuracy (Nustad, 1997).

Sampling Issues

Most sampling issues arise in obtaining measures of effectiveness of care and use of services. If the health plan has the necessary information in a complete and accurate computer database, then sampling may be relatively easy, but if this information is not available from a computer database, then the health plan must sample from physical medical records which can be far more costly and time consuming (Grimaldi, 1997). Many information systems may be inadequate for capturing and reporting the data required for HEDIS (Nustad, 1997).

Samples used for HEDIS are taken from the entire eligible population (HEDIS 3.0, 1998). Included in the denominator of HEDIS measures are those who have been continuously enrolled for a certain amount of time (usually one year), have an assigned primary care provider, and have complete coverage for the service being measured (Thompson, et al., 1998).

The NCQA has in recent years implemented consistent sampling procedures that are to be used to report data. These include new standardized sampling techniques for reporting some HEDIS measures, including standardizing how often a plan samples and who the plan includes as eligible members of the plan (Medical, 1997).

Simulation

Simulations were performed to find the sampling procedure that is most accurate in estimating population parameters. The simulations were performed with S-plus statistical software. Data was created using the binomial distribution at several probabilities of success. The populations created were each of size 100,000 with probabilities of .05, .10, .20, .30, .40, and .50. Because the populations were created with the binomial distribution, the actual population's probability of success vary slightly from the specified probability.

The simulations compare random, systematic, stratified, and stratified systematic sampling techniques. Random sampling involves the selection of n elements from a population of size N in which every possible sample of size n has an equal probability of being selected. Stratified sampling involves separating the population into nonoverlapping groups, or strata, and selecting random samples from each stratum. One example of possible strata is geography. For instance, if health plan utilization varies by a significant amount, strata might be urban vs. rural clinics, or each of a health plan's individual clinics might be considered strata. Systematic sampling involves randomly selecting a starting element from the first k elements and then every kth element thereafter. Stratified systematic sampling involves taking a systematic sample from each strata (Scheaffer, 1979).

Each of the populations were divided in half in order to create strata. Several of these divided populations were combined and random and systematic samples at various sample sizes were taken for comparative purposes. The results described in this paper will be for combinations of binomial populations at probabilities .05 and .50, .10 and .40, and .20 and .30. 5000 random and systematic samples of sizes 10, 30, 50, 100, and 200 from each of these combined populations were taken. From each of these samples the total number of successes were recorded. A success in this application represents a person who utilizes the measured medical service. Next, systematic samples were taken by creating a matrix of all possible sampling schemes (one for each starting point), and then randomly selecting the column in order to take each sample. In each of the individual populations, 5000 samples of sizes 5, 15, 25, 50, and 100 were taken using random sampling techniques. Samples from different strata were then combined to obtain stratified estimates of the population proportion, providing the stratified samples. Stratified systematic samples were also taken using systematic sampling techniques on individual strata and then combining the appropriate samples. In order to combine individual samples, an average of the sample proportion from the two sampled populations was taken.

Results

The appendix summarizes the results of the simulations. The results were obtained by comparing sample proportions from the various sampling techniques to the population proportions. Interest is focused on comparing the stratified sampling techniques with random and systematic sampling techniques. For each of the 5000 samples, the four techniques were analyzed to find the closest estimate to the population proportion. This analysis was divided into two sets of comparisons: random and systematic sampling vs. stratified sampling, and random and systematic sampling vs. stratified systematic sampling. This provides a clearer comparison between the three techniques. In the case of a tied estimate, the techniques were given a proportional fraction of the credit for having the best estimate. The two sets of analysis indicate the stratified systematic and stratified sampling techniques produce very similar results.

For the combination of the p = .05 and p =.50 populations, stratified sampling provides better results than either random or systematic sampling. When comparing random, systematic, and stratified sampling techniques, the simulations show that the stratified samples estimate the population proportion most accurately. The advantage is more pronounced for smaller sample sizes. For instance, for n=10, stratified sampling produced the best estimate in 32.1% of the samples, while systematic and random sampling only produced the best estimate in 26.7% and 27.2% of the samples respectively. For n=200, stratified sampling produced the best estimate in 27.4% of the samples, while systematic and random sampling produced the best estimate in 25.8% and 27.1% of the samples respectively. Stratified systematic sampling produces similar results. For instance, for n=10, stratified systematic sampling produced the best estimate in 31.6% of the samples, while systematic and random sampling only produced the best estimate in 26.3% and 27.1% of the samples respectively. For n=200, stratified systematic sampling produced the best estimate in 27.8% of the samples, while systematic and random sampling produced the best estimate in 25.9% and 27.3% of the samples respectively. For each sample size, both stratified estimates are better than random and systematic.

For the combination of the populations of p=.10 and p=.40, the stratified samples generally produce better results. The advantage of using stratified sampling is more significant for smaller sample sizes. The stratified and stratified systematic samples produced the best estimates of the population proportion a higher percentage of the time for n=10, 30, 50, and 100. In the case of n=200, the systematic sampling plan produced a better estimate in more of the samples when compared with either stratified or stratified systematic sampling plans.

For the combination of the populations of p=.20 and p=.30, there is no longer any clear advantage to the use of stratified sampling. In the case of stratified sampling, random sampling produces the best results for n = 10 and 30, systematic sampling produces the best results for n=100 and 200, while stratified sampling only produces the best results for n=50. In the case of stratified systematic sampling, random sampling produces the best results for n = 30, systematic sampling produces the best results for n=10, 50, and 100, while stratified systematic sampling only produces the best results for n=200. Therefore, there is no need to use stratified sampling strategies when the population proportions are close. Stratified sampling techniques can significantly improve estimates of population proportions if appropriate strata can be identified that differ by a moderate to large amount (a difference in proportions of .30 or larger). This result is more pronounced for smaller sample sizes (n = 100 or less).

Conclusion

The data collected through HEDIS and distributed in the Quality Compass are an important measurement of health plan performance. Some data is easily collected by participating plans through computer records, while other data must be collected through sampling physical patient records. New rules standardize the frequency of sampling and sampling rules (Medical, 1997). Currently, the only prescribed sampling plans for collecting the data are random and systematic methods. In simulations comparing random, systematic, stratified, and stratified systematic sampling methods, for several different probabilities and sample sizes, it is apparent that better results can be obtained using a stratified or stratified systematic sampling method if significantly different strata can be found. Stratified systematic sampling plans provide results that are similar to stratified sampling plans, but provide the additional advantage of an audit trail. The improvement in estimation is more pronounced for smaller sample sizes and larger differences in population proportions. In general, stratified systematic sampling should be used if sample proportions differ by .30 or more and sample sizes are 100 or less. Health plans interested in improving the accuracy of data presented to the NCQA should consider the value of a stratified systematic sampling plan.

References

De Lafuente, D. (1996), "Health Watchdog Issues Guidelines," Chicago Sun-Times, July 17, 1996, Pg. 69.

Grimaldi, P. (1997), "New HEDIS means more information about HMOs," Journal of Health Care Finance, NO. 4, Vol. 23, Pg. 40.

HEDIS 3.0 (1998), "Guidelines for Sampling and Calculations," Vol. 2, pp. 277-286.

Kristof, K. M. (1994), "A Plan To Help You Choose Among Health Plans," Los Angeles Times, October 16, 1994, D-6.

Medical Utilization Management (1997), "NCQA Facilitates, Standardizes Reporting," Author, No. 24, Vol. 25, pp. 1-2.

Nustad, C. J. (1997), "Measuring the Value of Health Care Dollars," Health Management Technology, No. 6, Vol. 18, Pg. 24.

Scheaffer, R. L., Mendenhall, W., Ott L. (1979), *Elementary Survey Sampling*, Duxberry Press, pp.22, 31, 59, 173-176.

Thompson, J. W., Bost, J., Ahmed, F., Ingalls, C. E., Sennett, C. (1998), "The NCQA's Quality Compass: Evaluating Managed Care In the United States," Health Affairs, January-February 1998, pp. 1-6.

Appendix

Comparing Random, Systematic, and Stratified Sampling

Population	n	Random	Systematic	Stratified
*p05-50	10	**26.70%	27.17%	**32.10%**
p05-50	30	25.74%	25.44%	**29.71%**
p05-50	50	25.83%	25.01%	**28.95%**
p05-50	100	24.67%	24.23%	**29.39%**
p05-50	200	25.81%	27.09%	**27.35%**
p10-40	10	27.65%	28.70%	**30.24%**
p10-40	30	27.15%	25.88%	**28.95%**
p10-40	50	25.62%	26.99%	**27.53%**
p10-40	100	24.47%	26.23%	**27.29%**
p10-40	200	25.00%	**27.96%**	27.12%
p20-30	10	**30.18%**	28.90%	27.93%
p20-30	30	**28.14%**	26.24%	27.27%
p20-30	50	26.95%	26.47%	**27.83%**
p20-30	100	25.04%	**27.94%**	26.20%
p20-30	200	26.60%	**27.00%**	25.66%

Comparing Random, Systematic, and Stratified systematic Sampling

Population	n	Random	Systematic	Stratified systematic
p05-50	10	26.30%	27.08%	**31.56%**
p05-50	30	25.45%	25.55%	**30.02%**
p05-50	50	26.04%	25.16%	**28.23%**
p05-50	100	24.62%	24.27%	**27.54%**
p05-50	200	25.94%	27.32%	**27.77%**
p10-40	10	27.49%	28.51%	**29.76%**
p10-40	30	27.63%	25.98%	**28.02%**
p10-40	50	25.65%	26.84%	**27.98%**
p10-40	100	24.25%	26.38%	**27.76%**
p10-40	200	24.88%	**28.07%**	27.79%
p20-30	10	29.89%	**28.96%**	28.38%
p20-30	30	**28.18%**	26.39%	27.22%
p20-30	50	26.83%	**26.85%**	26.16%
p20-30	100	24.95%	**27.95%**	26.07%
p20-30	200	26.56%	26.60%	**28.13%**

* p## - ## indicates which population is being sampled. For instance p05-50 represents the population obtained by combining the binomial populations at p = .05 and p = .50.

** Percentages indicate the percentage of 5000 sample proportions that were closest to the population proportion. (**Bold** percentages indicate that the corresponding sampling plan was most likely to produce the closest estimate of the population proportion.)